

Sonderdruck aus

Wolfgang Böttcher, Christiane Kerlen, Peter Maats, Oliver Schwab,  
Sonja Sheikh (DeGEval-Vorstand) (Hrsg.)

# Evaluation in Deutschland und Österreich

Stand und Entwicklungsperspektiven in den Arbeitsfeldern  
der DeGEval – Gesellschaft für Evaluation

Print-ISBN 978-3-8309-3149-2

E-Book-ISBN 978-3-8309-8149-7

© Waxmann Verlag GmbH, 2014  
Postfach 8603, 48046 Münster

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten. Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Bestellung per Fax: 0251 26504-26 oder  
telefonisch: 0251 26504-0;  
per Internet unter [www.waxmann.com/buch3149](http://www.waxmann.com/buch3149) oder  
per E-Mail: [order@waxmann.com](mailto:order@waxmann.com)

*Christian Erzberger, Udo Kelle*

## **Methodisch handeln im Evaluationsprozess**

Aus forschungsmethodischer Perspektive finden Evaluationen im Bereich der Praxis-, Anwendungs- und Auftragsforschung statt, greifen aber auch auf Befunde, Erfahrungen und Verfahren wissenschaftlicher Grundlagenforschung zurück. Die dort entwickelten Methoden, Strategien und Techniken müssen allerdings auf die spezifischen Bedürfnisse von Evaluation bezogen und daraufhin angepasst werden. Im Jahr 2010 wurde der AK „Forschungsmethoden in der Evaluation“ gegründet, um diese Probleme im Kontext der Fachgesellschaft DeGEval zu diskutieren. Im Gegensatz zu anderen Arbeitskreisen der DeGEval, die ein genuin inhaltliches Evaluationsfeld bearbeiten, ist die wesentliche Aufgabe unseres „Querschnittsarbeitskreises“ der offene Austausch über Evaluationsmethoden über einzelne inhaltliche Evaluationsfelder hinweg. Hierbei geht es um methodologische und methodische Probleme, die im Rahmen von Evaluationen auftreten sowie um Lösungsstrategien zur Bearbeitung dieser Probleme. Ein Schwerpunkt der Diskussionen liegt dabei auf besonderen Anforderungen und auch Beschränkungen, die auftragsgebundene Evaluationen mit sich bringen. Die methodische und methodologische Elaboration von Verfahren und Standards, wie sie etwa in DFG-Projekten an Universitäten entwickelt und umgesetzt werden, ist im Bereich der Evaluation aus praktischen Gründen oft nicht erreichbar, wäre auch teilweise kontraproduktiv, weil zeitliche Ressourcen beschränkt sind und Evaluationen unter dem Druck politischer und organisatorischer (d.h. betrieblicher und administrativer) Handlungserfordernisse und -erwartungen stehen. Sozialwissenschaftliche Forschung innerhalb eines Praxisfeldes (bzw. der Versuch, diese Praxis zu bewerten) ist einer „Operation am offenen Herzen“ vergleichbar – das bedeutet, dass sich Evaluation und die angewandten Methoden an den spezifischen Erfordernissen der jeweiligen Praxis orientieren müssen. Die Aufgabe, unter den spezifischen Bedingungen der Evaluationspraxis zuverlässige, valide und aussagekräftige Daten zu erheben und in sinnvoller Weise auszuwerten, ist äußerst anspruchsvoll. Evaluator/inn/en stehen permanent vor der Herausforderung, die in sozialwissenschaftlichen Fächern entwickelten Standards und Gütekriterien der Methodenentwicklung (wie Messgenauigkeit, Objektivität, inhaltliche Gültigkeit, Repräsentativität u.a.m.) mit den besonderen Bedingungen des Forschungsfeldes und den Anforderungen von Seiten der Auftraggebenden und anderer Stakeholder zu vereinbaren und in diesem Prozess unterschiedliche (akademische und administrativ-politische) Diskurswelten zueinander in Beziehung setzen. Dies setzt sowohl fachwissenschaftliche und methodische Kompetenz als auch eine

Offenheit bezüglich des evaluativen Vorgehens voraus – beides konstitutive Elemente des Handelns im Bereich der Evaluation.

„Offenheit“ bedeutet dabei, dass nicht nur über die klassischen Methoden zur Erhebung und Auswertung sozialwissenschaftlicher Daten reflektiert und diskutiert werden muss, sondern auch über besondere evaluative Verfahren wie etwa: entscheidungstheoretische oder beteiligungsorientierte Bewertungsverfahren. Unter „methodischer Kompetenz“ wiederum wird hier verstanden, dass – anders als in der Grundlagenforschung – in der auftragsbasierten Evaluation unter dem Handlungsdruck der Praxis oft die in den Fachwissenschaften bislang vorhandenen Standardlösungen aus Zeit- und Ressourcengründen nicht umgesetzt werden kann, sondern ökonomisch vertretbare und pragmatisch handhabbare Alternativen entwickelt werden müssen. Die methodische Strenge und Qualität und die methodologische Begründbarkeit solcher Alternativen, die gegenüber problematischen und unwissenschaftlichen Formen von Evaluationspraxis abgegrenzt werden müssen, ist ein ebenso wichtiges Anliegen für in der Evaluation Tätige und Auftraggebende, wie die Anwendung allgemein akzeptierter wissenschaftlich begründeter Qualitätskriterien im Handlungsfeld Evaluation. Hinzu kommt, dass auf dem sehr weiten und interdisziplinären Feld empirischer Forschungsmethoden und der Methodologie empirischer Sozial- und Verhaltensforschung in den letzten Jahrzehnten rapide Entwicklungen und bedeutsame Debatten stattgefunden haben. Diese betreffen einerseits die Entwicklung fortgeschrittener und teilweise recht komplexer Verfahren der Datenerhebung und Datenauswertung (vgl. etwa Diekmann 2004; Wolf, Best 2010) und andererseits konzeptuelle und grundsätzliche Fragen wie etwa das Verhältnis qualitativer und quantitativer Methoden (vgl. Kelle 2008, Kuckartz 2014). Viele dieser Entwicklungen sind für Evaluation als berufliche Praxis und wissenschaftlich begründete Tätigkeit von unmittelbarer Relevanz: dies gilt etwa für die zunehmende Verfügbarkeit von EDV-gestützten, statistisch anspruchsvollen, aber oft leicht zu implementierenden Datenanalysetechniken, deren Anwendungsvoraussetzungen nichtdestotrotz sehr komplex sein können. Das betrifft weiterhin Debatten über „Hierarchien der Evidenz“ und „evidence based practice“ u.a.m (vgl. bspw. Grinnell, Unrau 2010, Hammersley 2013). Von diesen Entwicklungen angeregt, oder auch unabhängig davon, finden in wissenschaftlichen Einzeldisziplinen immer wieder intensive Debatten über wissenschaftlich begründete Standards für gute und vertretbare Forschung statt. Der AK „Forschungsmethoden in der Evaluation“ greift diese Debatten und ihre Ergebnisse auf und diskutiert deren Relevanz für die Evaluationspraxis.

## 1. Ziele des Arbeitskreises

In dem Arbeitskreis wird der aktuelle Stand der Methodendiskussion, der Methodenentwicklung und der Debatte um Bewertung und Bewertungsverfahren in relevanten Fachdisziplinen (Psychologie, Statistik, Soziologie, Ökonomie, Medizin, Pflegewissenschaften u.a.m.), insoweit dieser Stand für Evaluation von Relevanz ist, diskutiert, um auf diese Weise unterschiedliche, für Evaluation relevante Fachdisziplinen in Austausch miteinander zu bringen. Im Zentrum der Aufmerksamkeit stehen dabei Gütestandards und Qualitätskriterien von Evaluationsmethoden. Im Sinne der aus der evidenzbasierten Medizin bekannten „Hierarchie der Evidenz“ sollen dabei *best practi-*

ces und akzeptable *good practices* identifiziert werden und die Geltungsreichweite von Aussagen, die mit Verfahren unterschiedlicher methodischer Strenge und Komplexität formuliert werden können, vergleichend bestimmt werden. Damit soll die Bandbreite wissenschaftlich vertretbarer Evaluationsmethoden in zwei Richtungen abgegrenzt werden: einerseits gegenüber zeit- und personalintensiven, hochkomplexen methodischen Verfahren, die in der wissenschaftlichen Grundlagenforschung einen Platz haben und andererseits gegenüber „Folkloremethoden“ der Praxis, denen methodische Kontrolle weitgehend fehlt. Für in der Evaluationsforschung Tätige ebenso wie für Auftraggebende und an Evaluationsergebnissen Interessierte sollen auf diese Weise Brauchbarkeit und Erkenntnisgrenzen unterschiedlicher methodischer Ansätze verdeutlicht werden. Hierbei werden praxisferne Debatten um abstrakte methodische und methodologische Konzepte „am grünen Tisch“ vermieden, grundsätzliche methodische Fragen und Probleme vielmehr beständig in Referaten und Plenums- und Kleingruppendiskussionen auf die Erfordernisse der Evaluationspraxis bezogen.

## 2. Methodische Problembereiche

Im Folgenden werden wir einige der wesentlichen methodischen Problemfelder in der Evaluation skizzieren, in denen der AK-Methoden aktiv ist:

### a) Qualitative und quantitative Daten und Methoden

Die starke Trennung zwischen der qualitativen und der quantitativen Methodentradition in der empirischen Sozialforschung erweist sich gerade im Bereich der Evaluation als schwerwiegendes Problem, weil qualitative und quantitative Methoden für jeweils unterschiedliche Ziele einsetzbar sind, die im Rahmen von Evaluationsdesigns oft gemeinsam verfolgt werden müssen: dies ist einerseits die Notwendigkeit, objektiv messbare und für bestimmte Gruppen verallgemeinerbare Ergebnisse, etwa über die Wirkungen eines bestimmten Programms, zu produzieren. Hierzu ist der Einsatz quantitativer Verfahren unabdingbar. Andererseits verfolgen Evaluationen aber auch häufig das Ziel, bislang unbekannte Prozesse und Strukturen und die ihnen zugrunde liegende Wirkkräfte zu identifizieren. Die Exploration bislang unbekannter Sachverhalte und Zusammenhänge ist eine klassische Domäne qualitativer, interpretativer Methoden der Sozialforschung. Nun generieren beide Methodentraditionen bei der Erfassung sozialer Wirklichkeit oft spezifische „blinde Flecken“, die erst durch eine Kombination qualitativer und quantitativer Daten und Methoden entdeckt und ausgeglichen werden können. Die Einsicht, dass qualitative und quantitative Forschung jeweils spezifische Schwächen und Begrenzungen aufweisen, die durch die Stärken der jeweils anderen Methodentradition ausgeglichen werden können, spielt eine große Rolle für die in den letzten beiden Jahrzehnten aus der US-amerikanischen Bildungsforschung entstandenen „Mixed methods“ Bewegung (Tashakkori, Teddlie 2010), die in den Diskussionen in der Evaluation einen prominenten Platz einnimmt. Dabei hat die Frage nach der Bedeutung methodenübergreifender Gütekriterien und Qualitätsstandards eine zentrale Bedeutung: Die Etablierung von allgemein akzeptierten Standards für die Wissenschaftlichkeit von Evaluationen ist ein zentraler Schritt hin zur Professionalisierung von Evaluation sowohl im Sinne der internen Diskussion und

Verständigung innerhalb der Profession als auch im Hinblick auf die Außenwirkung gegenüber (potenziellen) Auftraggebern und der Öffentlichkeit.

#### b) Relevanz von Evaluationsergebnissen

Im Unterschied zur Grundlagenforschung sollen Ergebnisse von Evaluationen in der Regel dazu dienen, Handlungspraxis direkt zu bewerten und anzuleiten. Dieser Fokus bringt besondere Herausforderungen mit sich, weil Fragen nach organisatorischen und institutionellen Rahmenbedingungen, unter denen Evaluationen stattfinden, hier besondere Aufmerksamkeit verdienen: Evaluation darf einerseits die Praxis, auf die sie sich bezieht, nicht zu sehr behindern, andererseits aber dürfen deren Erfordernisse auch die Evaluation nicht in der Erhebung valider Daten und der Ermittlung von Ergebnissen einschränken. Um die Balance zwischen professionellen Standards einerseits und den Bedarfen des Evaluationsfeldes andererseits zu halten, ist es unumgänglich, relevante Programmakteurinnen und -akteure in die Untersuchung einzubeziehen, welche die Implementation von Evaluationen beratend begleiten, ohne dabei deren Ziele zu gefährden. Die Aufgabe der Evaluatorinnen und Evaluatoren besteht dann in der Aushandlung von für die Beteiligten tragbaren und professionell verantwortbaren Kompromissen beim Einsatz unterschiedlicher Methoden. In diesen Zusammenhang gehören auch die Belange des Datenschutzes, dessen Auflagen gerade in sozial sensiblen Evaluationsfeldern sinnvolle Evaluationen oft sehr erschweren.

#### c) Methodische Qualität von Evaluation

Die Praxisgebundenheit von Evaluation hat nicht nur Folgen für die Implementierung der Untersuchung in die Programmstrukturen, sie müssen nämlich häufig in den täglichen Prozessablauf einer Institution eingebaut werden. Dies kann auch die angewandten Methoden selbst mehr oder weniger stark beeinflussen: Zeitliche und finanzielle Restriktionen erschweren bspw. oftmals den Einsatz aufwändiger Verfahren qualitativer und quantitativer Datenerhebung und -auswertung. Bei der Durchführung qualitativer Interviews sind dann zum Beispiel Transkriptionen der aufgezeichneten Gespräche schwierig zu realisieren; oder bei quantitativen Erhebungen überschreitet die nach stichprobentheoretischen Kriterien erforderliche Größe des Samples den vom Auftraggeber gesteckten Budgetrahmen. Die Rolle wissenschaftlicher Kriterien und Standards in Evaluationsprozessen und die Bedeutung unterschiedlicher Wissenschaftsverständnisse in diesem Kontext ist deshalb ein zentrales Diskussionsfeld für die Evaluation. Dies betrifft auch Fragen nach der wissenschaftlichen Dignität und Seriösität von unterschiedlichen Methoden der Datensammlung und -auswertung, die in der Evaluationspraxis eingesetzt werden.

#### d) Beziehungen zwischen Evaluator/inn/en und Auftraggebern

Bei Evaluationen nehmen beteiligte und betroffene Akteure mit divergierenden Interessen und unterschiedlichen Machtressourcen auf die Prozesse der Zielformulierung, Instrumentenkonstruktion, Datenerhebung, Datenanalyse und Dateninterpretation Einfluss. Interessen von unterschiedlichen Stakeholdern können dabei mit professionellen Standards von Evaluation und Evaluatoren in Konflikt geraten, aber auch Auftraggebende können sich in unterschiedlichen Dilemmata zwischen organisationsbezogenen Anforderungen, ökonomisch-politischen Sachzwängen und professionel-

len Ansprüchen wiederfinden. Diese Probleme müssen im Rahmen von Evaluationen reflektiert werden, und es müssen forschungspraktikable Lösungen gefunden werden, die einerseits die jeweiligen Interessen der verschiedenen Stakeholder berücksichtigen und bei denen andererseits Kriterien methodisch professioneller Evaluation nicht verletzt werden.

#### e) Vermittlung und Dissemination von Ergebnissen

Im universitären oder außeruniversitären Wissenschaftsbetrieb treffen die Darstellung von Ergebnissen und die Beschreibung des methodischen Vorgehens normalerweise auf ein Fachpublikum, das mit wissenschaftlich komplexen elaborierten Codes vertraut ist. In Evaluationen ist dies dahingegen nicht immer der Fall. Hier müssen Aussagen und Fachtermini in einem für alle Beteiligten verständlichen Code übermittelt werden und dabei die Komplexität der Darstellung (bspw. der verwendeten Methoden) an vorhandene Wissensbestände angepasst werden, ohne dass inhaltliche Aussagen verfälscht oder unvollständig wiedergegeben werden.

### 3. Entwicklungsperspektiven

Die oben skizzierten methodischen Herausforderungen geben in aller Kürze die Arbeitsfelder wieder, die den Rahmen für die Diskussionen des AK-Methoden abgeben. In unserem Querschnittsarbeitskreis werden diese Probleme an unterschiedlichen inhaltlichen Beispielen aus teilweise sehr unterschiedlichen Evaluationsfeldern bearbeitet. Dabei bemühen wir uns stets um ein lösungsorientiertes Vorgehen, indem etwa die Vortragenden dazu motiviert werden, in ihren Präsentationen Beispiele für mögliche Lösungen anzubieten. Auch in den an die Vorträge sich anschließenden Gruppenarbeitsphasen werden lösungsorientierte Aufgabenstellungen formuliert. Diese Einstellung hat auch die Auswahl der Themen für die bisherigen Frühjahrstagungen stark beeinflusst.

Wie alle anderen Arbeitskreise auch führt der AK „Forschungsmethoden in der Evaluation“ Sessions auf den Jahrestagungen der DeGEval sowie jährlich eine eigene Frühjahrstagung durch. Im Zentrum standen in der Vergangenheit dabei immer praktische, innovative methodische Ansätze von Evaluationen und methodisch-methodologische Grundsatzfragen in ihrem Bezug zur Evaluationspraxis. Die jeweils inhaltliche Thematik hatte dabei lediglich die Funktion eines „Trägers“, auf der die methodischen Verfahren erläutert wurden. An die Vorträge schließen sich Diskussionen an, in denen methodische und methodologische Debatten und Konzepte auf praktische Fragestellungen, Arbeitsfelder und Erfahrungen bezogen werden. Diese Diskussionen wurden verschriftlicht und als Ergebnisprotokolle den Mitgliedern der DeGEval bzw. den am AK-Methoden Interessierten zur Verfügung gestellt. Das Themenspektrum der Frühjahrstagungen orientiert sich dabei an dem ursprünglichen „Mission Statement“ des Arbeitskreises aus der Gründungsphase und an einer Befragung der Mitglieder des AK-Methoden bezüglich der inhaltlichen Ausrichtung des Arbeitsspektrums. Die Tagungen sind jeweils unterteilt in Präsentationen einerseits und Kleingruppenarbeit andererseits, in der die Beiträge diskutiert werden. In der Regel gibt es vier Vorträge,

zwei, die sich eher mit methodischen bzw. methodologischen Grundsatzproblemen beschäftigen und zwei, die eher praktisch ausgerichtet sind.

Die folgenden Themen wurden in der Vergangenheit bearbeitet:

- 2011 „Wirkungsanalyse – qualitative und quantitative Ansätze“
- 2012 „Best Practice‘ zwischen wissenschaftlichen Standards und politischer Einflussnahme“
- 2013 „Zwischen wissenschaftlichen Gütekriterien und den Erfordernissen der Evaluationspraxis I – die quantitative Seite“
- 2014 „Zwischen wissenschaftlichen Gütekriterien und den Erfordernissen der Evaluationspraxis II – die qualitative Seite“.

Relevante Themen für den AK wurden bereits in dem Gründungspapier des AK aufgelistet; auf einer der ersten Frühjahrstagungen wurde diese Liste durch ein Meinungsbild unter den AK-Mitgliedern ergänzt:

- Forschungsdesigns in der Evaluation
- Gütekriterien empirischer Forschung
- Sampling in der Evaluation
- Verfahren standardisierter Datenerhebung
- Qualitative Interview- und Beobachtungsverfahren
- Schärfung von Begrifflichkeiten
- Statistische Methoden in der Evaluation
- Pragmatismus versus „Gold-Standard“
- Übersetzung von Ergebnissen an Laien
- Partizipative Verfahren und ihre Grenzen

Auf den letzten Jahrestagungen ist es dabei gelungen, eine lebendige und fruchtbare Debatte zwischen Evaluatoren und Evaluatorinnen, Auftraggebenden und zu Evaluationsmethoden arbeitenden Wissenschaftler/inne/n über methodische Fragen und Probleme zu initiieren und dabei sowohl die Frage der Umsetzbarkeit von professionellen Evaluationsmethoden in der Evaluationspraxis als auch die Bedeutung von methodischen Gütekriterien und Qualitätsstandards angemessen zu berücksichtigen.

## Literatur

- Diekmann, Andreas (2004) (Hg.), Methoden der Sozialforschung, Sonderheft 44/2004 der Kölner Zeitschrift für Soziologie und Sozialpsychologie. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Grinnell, Richard M. Jr.; Unrau, Yvonne A. (2013): Social Work Research and Evaluation: Foundations of Evidence-Based Practice. Oxford: Oxford University Press.
- Hammersley, Martyn (2013): The Myth of Research-based Policy and Practice. London: Sage.
- Kelle, Udo (2008): Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung. Theoretische Grundlagen und methodologische Konzepte. Wiesbaden: Verlag für Sozialwissenschaften.

- Kuckartz, Udo (2014): *Mixed Methods. Methodologie, Forschungsdesigns und Analyseverfahren*. Wiesbaden: VS Verlag.
- Tashakkori, Abbas; Teddlie, Charles (2010): *Sage Handbook of Mixed Methods in Social and Behavioral Research* (2. Auflage). Thousand Oaks: Sage.
- Wolf, Christof; Best, Henning (2010): *Handbuch der sozialwissenschaftlichen Datenanalyse*. Wiesbaden: VS/Springer.